

INTISARI

SELEKSI FITUR MENGGUNAKAN *ANT COLONY OPTIMIZATION* PADA IDENTIFIKASI *SPAM*

Oleh

Harry Pribadi Fitrian

20/466410/PPA/05976

Email menyediakan layanan yang cepat, gratis, serta memiliki kemampuan dapat mengirim pesan ke banyak penerima dalam waktu singkat. Terdapat pihak tertentu yang menyalahgunakannya dengan mengirim *email* berisi promosi produk dan hal-hal yang tidak penting ke pengguna *email* lainnya. *Email* tersebut disebut *email SPAM* (*Stupid Pointless Annoying Message*).

Banyak kerugian yang disebabkan oleh *email SPAM* maka diperlukannya sistem yang dapat mendeteksi *email SPAM* yang akurat. Masalah utama dalam hal mendeteksi *email SPAM* yaitu dimensi klasifikasi teks yang lebih tinggi dari ruang fitur, yang sering terjadi dalam teks yang memiliki puluhan ribu fitur. Kebanyakan dari fitur ini tidak relevan dan tidak berguna untuk klasifikasi teks bahkan dapat mengurangi tingkat akurasi.

Penelitian ini bertujuan untuk menyeleksi fitur yang paling tepat untuk mengurangi waktu komputasi dan mengetahui waktu komputasi dari kinerja seleksi fitur menggunakan *ant colony optimization* untuk aplikasi *filtering email* yang memanfaatkan metode *Naive Bayes Classifier*. Banyak metode yang diusulkan untuk seleksi fitur, algoritma optimasi berbasis populasi seperti *Ant Colony Optimization* (ACO) telah menarik banyak perhatian. Hasil eksperimen menunjukkan *Ant Colony Optimization* dapat menyeleksi fitur hingga 92,9%. Hasil akurasi dengan menggunakan seleksi fitur *Ant Colony Optimization* adalah 96,89% dan tanpa menggunakan seleksi fitur adalah 93,44 %.

Kata Kunci: *Email, SPAM, Naive Bayes Classifier, Ant Colony Optimization*

ABSTRACT

FEATURE SELECTION USING ANT COLONY OPTIMIZATION ON SPAM IDENTIFICATION

by

HARRY PRIBADI FITRIAN

20/466410/PPA/05976

Email provides fast, free service, and has the ability to send messages to many recipients in a short time. There are certain parties who abuse it by sending emails containing product or service promotions and things that are not important to other email users. These emails are called SPAM (Stupid Pointless Annoying Message) emails.

Many losses are caused by SPAM emails, so we need a system that can detect SPAM emails accurately. The main problem with detecting SPAM emails is the higher text classification dimension of the feature space, which is often the case in text that has tens of thousands of features. Most of these features are irrelevant and useless for text classification and can even reduce the level of accuracy.

This research aims to select the most appropriate features to reduce computation time and determine the computational time of feature selection performance using ant colony optimization for email filtering applications that utilize the Naive Bayes Classifier method. Among the many proposed methods for feature selection, population-based optimization algorithms such as Ant Colony Optimization (ACO) have attracted much attention. The experimental results show that Ant Colony Optimization can select features up to 92.9%. The result of accuracy using Ant Colony Optimization as feature selection is 96.89% and without using feature selection is 93.44%.

Key Word: *Email, SPAM, Naive Bayes Classifier, Ant Colony Optimization*